

MC, the DB and SAM

Rick St. Denis – Glasgow

- The DB
- Datasets in SAM: what is there now and what is planned
- What will be done

DB access in MC

- Look at what is accessed at
 - Begin Job
 - Begin Run
 - Event
- Use Action on “Name Action” and setenv DEBUG 3 to see what an MC job constructed using web instructions does.

Accesses:

- Begin Job: CprClusterModule,
CentralStripClusterModule,
SiClusteringModule
- **Begin Run: Calibration Manager**
- Event: SimulationControlMod

What they Access

- CprClusteringModule:
CALDigitToGev3,CALL1Peds3,CHASlewing,PH
ASlewing,WHASlewing,PPRLinERResponse ->
All Fail! Good thing: wrong constants.
- CentralStripClusterModule: Creates Managers
- SiClusteringModule: SiDpsParms (3932 rows
from run 138623 version 1)
- Calibration Manager: PROD_PHYSICS_CDF for
run 146 v 7 (fails)
- Calorimetry Module: (Creates Managers)
 - PesDtoEModule (Prints “No calib in MC”)
- SimulationControlMod (requests managers)

Problems with this

- Calibrations accessed in begin job BEFORE the manager can decide what you should have
- Mixed policy of allowing code to decide what to do and allowing calibration management.
- TCL contains information on Silicon calibration run to use.

Suggested solution

- Create valid sets (using tools of used sets) by analysis module: CALDigitToGev3 etc should be in one.
- MC request managers create a used set for the MC run.
- MC validation should include check using the tools action on “Name Action” and setenv DEBUG 3
- Most valid sets should be made once and used many times. Dead channel handling used in reconstruction.
- Bookkeeping: see sam stuff.

Suggestion Solution

- MC reps from physics groups manage the calibrations. Naturally addresses
 - How to handle lots of parameters, a physics question.
 - Do you want dead ladders? How many CDF detector configurations do you analyze?
 - Motivating coders to get the code sorted.

SAM: DFC and MC

- Everything in DFC is automatically transferred to sam
- mtestc contains lots of datasets
- Filename is mc....test.
- Run numbers distinguish processes
- Look for data in current MC pages and in SAM.

MC requested 4.5.0

Request id	-	Run number	85899	Phys. Grp.	Top/Ewk	Gen type	Herwig
Events	246k	Xsect (pb)	4.652	Lum. (pb^{-1})	52880	Req. St.	done
User	Yacintus	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready
Physics type	tbar	Files	82	Events per file	3000	View tcl	

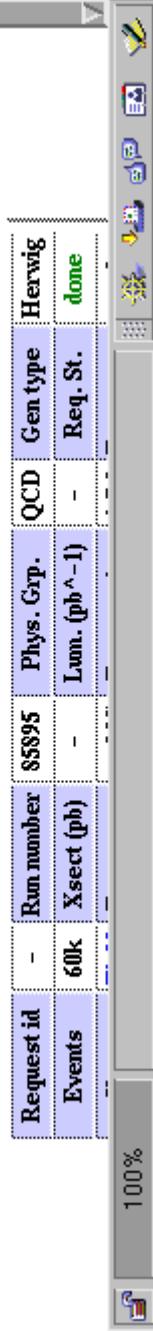
Description text
inclusive tbar

Request id	-	Run number	85893	Phys. Grp.	QCD	Gen type	Herwig
Events	60k	Xsect (pb)	-	Lum. (pb^{-1})	-	Req. St.	done
User	Field	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready
Physics type	QCD	Files	12	Events per file	5000	View tcl	

Description text
QCD dijets, pt>5 GeV

Request id	-	Run number	85894	Phys. Grp.	QCD	Gen type	Herwig
Events	60k	Xsect (pb)	-	Lum. (pb^{-1})	-	Req. St.	done
User	Field	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready
Physics type	QCD	Files	12	Events per file	5000	View tcl	

Description text
QCD dijets, pt>12 GeV



MC requested 4.5.0

Request id	-	Run number	85899	Phys. Gtp.	Top/Ewk	Gen type	Hewwig
Events	246k	Xsect (pb)	4.652	Lum. (pb^-1)	52880	Req. St.	done
User	Vaitulis	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready
Physics type	tthar	Files	\$2	Events per file	3000	View tcl	

Description text
inclusive tthar

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: http://cdfdb.fnal.gov/sam_data_browsing/ProjectDefinitions.html What's Related

DBStat Farms Processing System RealTune Yahoo! Radio LSF CBrow DBBrowser trightbit RStD D

SAM Dataset Definitions

[Clear Form](#)

[Instructions](#)

Note: DatasetDefinitions are synonymous with their old names, ProjectDefinitions.

Hint: Use percentkey (%) for a wildcard. Names are case-sensitive.

Dataset Definition Name:	<input type="text" value="2_Interest"/>
Persons First Name:	<input type="text"/>
Persons Last Name:	<input type="text"/>
Username:	<input type="text"/>
Physics Work Group:	<input type="text" value="2_"/>
Started Before: (dd-mon-yyyy) Ex: 06-JUN-1999	<input type="text"/>
Started On or After: (dd-mon-yyyy) Ex: 06-JUN-1999	<input type="text"/>
Sort Order:	<input type="button" value="Create Date (Descending)"/>
Maximum rows of output:	<input type="text" value="1000"/> displaying <input type="text" value="100"/> per page
Select Output Format:	<input type="button" value="HTML"/>
Fetch Matching Dataset Definitions:	<input type="button" value="Run"/>

Query Tips

Construct your own SQL query: Build

100%

Dataset Definition Name:	<input type="text" value="Intests"/>
Persons First Name:	<input type="text"/>
Persons Last Name:	<input type="text"/>
Username:	<input type="text"/>
Physics Work Group:	<input type="text" value="?"/>
Started Before: (dd-mon-yyYY) Ex: 06-JUN-1999	<input type="text"/>
Started On or After: (dd-mon-yyYY) Ex: 06-JUN-1999	<input type="text"/>
Sort Order:	<input type="button" value="Create Date (Descending)"/>
Maximum rows of output:	<input type="text" value="100"/> per page
Select Output Format:	<input type="button" value="HTML"/>
Fetch Matching Dataset Definitions:	<input type="button" value="Run"/>

IAM – Dataset Definitions

Name	Create Date	Person	User Name	Work Group	Description
mtestc-top	29-JUL-02	St. Denis, Richard	stdenis	top	inclusive ttbar test
mtestc	03-JUN-02	St. Denis, Richard	stdenis	filecatalog	run ii commissioning, monte carlo test

rows 1 to 2 of the Total 2 found.

Note the mtestc is same as in dfc:

Work Group is filecatalog

I will show you how to make the mtestc-top dataset
You have freedom to clone and restrict
datasets as you wish

Test	Date	Author	Description	Category
itestc-top-175-462	31-JUL-02	St. Denis,Richard	stdenis	top
itestc-gamma-jet-pt12-eta11	31-JUL-02	St. Denis,Richard	stdenis	top
itestc-bbbar-lepton-pr10-is015	31-JUL-02	St. Denis,Richard	stdenis	b physics
itestc-qcd-dijet-pt20	31-JUL-02	St. Denis,Richard	stdenis	qcd
itestc-z-gamma-tautau	31-JUL-02	St. Denis,Richard	stdenis	electroweak
itestc-incl-ww	31-JUL-02	St. Denis,Richard	stdenis	electroweak
itestc-w2emu-453	31-JUL-02	St. Denis,Richard	stdenis	electroweak
itestc-drellyan-z2ee-453-v2	31-JUL-02	St. Denis,Richard	stdenis	electroweak
itestc-drellyan-z2ee-453	31-JUL-02	St. Denis,Richard	stdenis	electroweak
itestc-ex-zimumu-eta5-pt8-nomunu	31-JUL-02	St. Denis,Richard	stdenis	exotics

testc-ex-zmmumu-eta5-pt8-mumin	31-JUL-02	St. Denis,Richard	stdenis	exotics	drell yan z->mu mu, eta <5, pt>8 gev muon miniskirt turned on.
testc-ex-zmmumu-eta3-pt18-set2	31-JUL-02	St. Denis,Richard	stdenis	exotics	drell yan z->mu mu, eta <3, pt>18 gev muon miniskirt turned off.
testc-ex-zmmumu-eta3-pt18-set1	31-JUL-02	St. Denis,Richard	stdenis	exotics	drell yan z->mu mu, eta <3, pt>18 gev muon miniskirt turned on.
testc-ex-2gamer22-eta3	31-JUL-02	St. Denis,Richard	stdenis	exotics	cm energy=2 tev, pythia default pdf at hepg stage, applied a filter requiring 1 photon with et>22 gev and eta <3.0, this has a filter rate of 0.251, only events passing are simulated and written out. 0 all included subprocesses i36000 122507 i3.647 e-05 (mb) 14 f + fbar -> g + gamma i4260 12862 i 4.342e-06 (mb) 29 f + g -> f + gamma i31728 109619 i3.212e-05 (mb) 115 g + g -> g + gamma i 12 26 i 1.361e-08 (mb)
testc-ex-2gamer10-eta3	31-JUL-02	St. Denis,Richard	stdenis	exotics	cm energy=2 tev, pythia default pdf at hepg stage, applied a filter requiring 2 photons with et>10 gev and eta <3.0, this has a filter rate of 0.124, only events passing are simulated and written out. 0 all included subprocesses i36000 119424 i7.421 e-07 (mb) 18 f + fbar -> gamma + gamma i18491 71857 i 3.820e-07 (mb) 114 g + g -> gamma + gamma i 17509 47567 i 3.600e-07 (mb)
testc-qcd-dijet-pt150	31-JUL-02	St. Denis,Richard	stdenis	qcd	qcd dijets, pt>150 gev
testc-qcd-dijet-pt70	31-JUL-02	St. Denis,Richard	stdenis	qcd	qcd dijets, pt>70 gev
testc-qcd-dijet-pt30	31-JUL-02	St. Denis,Richard	stdenis	qcd	qcd dijets, pt>30 gev
testc-qcd-dijet-pt12	31-JUL-02	St. Denis,Richard	stdenis	qcd	qcd dijets, pt>12 gev
testc-qcd-dijet-pt5	31-JUL-02	St. Denis,Richard	stdenis	qcd	qcd dijets, pt>5 gev

Name	CreateDate	Person	UserName	WorkGroup	Description
testc-drellyan-zmumu-430pre1 and2	31-JUL-02	St. Denis,Richard	stdenis	electroweak	clone of 349. Pythia drell-yan z to ee run number 85974; runmc based on 4.3.0pre1 using cvs tag v4_3_0pre2 for mcproduction.
testc-drellyan-zee-430pre1 and2	31-JUL-02	St. Denis,Richard	stdenis	electroweak	pythia drell-yan z to ee run number 85974; runmc based on 4.3.0pre1 using cvs tag v4_3_0pre2 for mcproduction.
testc-wwgeneric-430pre1 and2	31-JUL-02	St. Denis,Richard	stdenis	electroweak	pythia ww generic run number 85975; runmc based on 4.3.0pre1 using cvs tag v4_3_0pre2 for mcproduction.
testc-z2mmumu-430pre1 and2	31-JUL-02	St. Denis,Richard	stdenis	electroweak	pythia z to mmum on-shell run number 85976; runmc based on 4.3.0pre1 using cvs tag v4_3_0pre2 for mcproduction.
testc-z2eee-430pre1 and2	31-JUL-02	St. Denis,Richard	stdenis	electroweak	* pythia z to ee on-shell run number 85977; runmc based on 4.3.0pre1 using cvs tag v4_3_0pre2 for mcproduction.
testc-w2emu-430pre1 and2	31-JUL-02	St. Denis,Richard	stdenis	electroweak	pythia w to munu run number 85978; runmc based on 4.3.0pre1 using cvs tag v4_3_0pre2 for mcproduction.
testc-w2emu-420	31-JUL-02	St. Denis,Richard	stdenis	electroweak	* pythia w to emu run number 85979; runmc was patched extensively. cvs tag usewithv4_2_0 for mcproduction.
testc-top	29-JUL-02	St. Denis,Richard	stdenis	top	inclusive ttbar test
testc	03-JUN-02	St. Denis,Richard	stdenis	filecatalog	run ii commissioning, monte carlo test

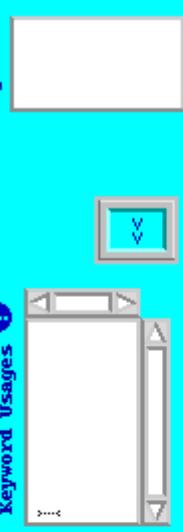
SMW Dataset Definition Details

Dataset definition has already been used to create actual datasets, so you cannot edit it.
You may, however, add keywords to help you find it later.
For more information, see the datasets created using this definition.

Definition Name: mcstest

Keywords

Keyword Usages [?](#)



Definition Id: 257

Create Date: 06/03/2002 16:05:11

Work Group: filecatalog

Username: sdenis

Description: run li commissioning, monte carlo test

Dimension Query
FILE_NAME mcstest

Translate Constraints Results

[Clone](#) [Create Dataset](#)

```
Total File Count: 1670 Total Event Count: 8026945 Rvg File Size: 45449
mc014ffa.0001test
mc014ffa.0002test
mc014ffa.0003test
mc014ffa.0004test
mc014ffa.0007test
mc014ffa.0008test
mc014ffa.0009test
mc014ffa.000atest
mc014ffa.000btest
```

100%

Definition Name: mtestc

Definition Id: 257

Create Date: 06/03/2002 16:05:11

Work Group: filecatalog

Username: stdenis

Description: run ii commissioning, monte carlo test

Dimension Query
FILE_NAME mtestc

Translate Constraints Results

[Clone](#)

[Create Dataset](#)

```
Total File count: 1670 Total Event Count: 8026945 Avg File size: 45449
mc014f6a.0001test
mc014f6a.0002test
mc014f6a.0003test
mc014f6a.0004test
mc014f6a.0007test
mc014f6a.0008test
mc014f6a.0009test
mc014f6a.000atest
mc014f6a.000btest
```

Total File count: 1670 Total Event Count: 8026945

Definition Name:	<input type="text"/> iclone-mt-test	Keywords
Definition Id:	<input type="text"/> ?	Keyword Usages
Create Date:	<input type="text"/> ?	< auto-assigned >
Work Group:	<input type="text"/> ?	< auto-assigned >
Username:	<input type="text"/> ?	<input type="checkbox"/> all <input type="checkbox"/> none
Description:	<input type="text"/> ?	<input type="button"/> clear
Dimension Query	<input type="text"/> ? FILE_NAME mt-test	<input type="button"/> Save
<p>If you enter both a detailed query above and constraints below, they will be combined to make one query. or, use the Combine Query option to see the query before translating it into a set of files or saving it.</p>		
Operator	Dimension	constraint value
and	<input type="checkbox"/>	<input type="text"/> RUN_NUMBER
and	<input type="checkbox"/>	<input type="text"/>
and	<input type="checkbox"/>	<input type="text"/>
<input type="button"/> Translate Constraints		

SRM Dataset Definition Details

Dataset Name:	<input type="text" value="Clone-Mttest0"/>
Definition Id:	<input type="text" value="< auto-assigned >"/>
Create Date:	<input type="text" value="< auto-assigned >"/>
Work Group:	<input type="text" value=""/>
Username:	<input type="text" value=""/>
Description:	<input type="text" value="Clone of 257. run ii commissioning, monte carlo te"/>

If you enter both a detailed query above and constraints below, they will be combined to make one query.

Or, use the [Combine Query](#) option to see the query before translating it into a set of files or saving it

operator Dimension

and

RUN_NUMBER

85899

and

and

MC requested 4.5.0

Request id	-	Run number	85899	Phys. GrP.	Top/Ewk	Gen type	Herwig
Events	246k	Xsect (pb)	4.652	Lum. (pb^-1)	52880	Req. St.	done
User	Vaitulis	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready
Physics type	tthar	Files	82	Events per file	3000		View tcl

Description text

inclusive tbar

Definition Name:	<input type="text" value="Clone-mctestc"/>	Keywords	<input type="text"/>
Definition Id:	<input type="text" value="2"/>	Operator	<input type="text" value="< auto-assigned >"/>
Create Date:	<input type="text" value="2"/>	Dimension	<input type="text" value="< auto-assigned >"/>
Work Group:	<input type="text" value="2"/>	Value	<input type="text" value="none"/>
Username:	<input type="text" value="2"/>	Constraint	<input type="text" value="all"/>
Description:	<input type="text" value="Clone of 257. run ii commissioning, monte carlo te"/>	Value	<input type="text" value="clear"/>
Dimension query		<input type="text" value="FILE_NAME mc%test and RUN_NUMBER 85899"/>	
<p>If you enter both a detailed query above and constraints below, they will be combined to make one query.</p> <p>Or, use the Combine Query option to see the query before translating it into a set of files or saving it.</p>			
operator	constraint	value	constraint
<input type="text" value="and"/>	<input type="text" value="I"/>	<input type="text" value="and"/>	<input type="text" value="I"/>
<input type="text" value="and"/>	<input type="text" value="I"/>	<input type="text" value="and"/>	<input type="text" value="I"/>
<input type="text" value="and"/>	<input type="text" value="I"/>	<input type="text" value="and"/>	<input type="text" value="I"/>
		Save	Create Dataset
		Translate Constraints	
<pre>Total File Count: 82 Total Event Count: 246000 Avg File Size: 413883 mc014f8b.004etest mc014f8b.0043test mc014f8b.0045test mc014f8b.003atest mc014f8b.003ftest mc014f8b.0040test mc014f8b.003dtest mc014f8b.0041test mc014f8b.0044test</pre>			

Dimension query

FILE_NAME mc%test and RUN_NUMBER 85899

Total File count: 82 Total Event count: 246000 Avg File size: 413383
mc014f8b_004etest
mc014f8b_0043test
mc014f8b_0045test
mc014f8b_003etest
mc014f8b_003ftest
mc014f8b_0040test
mc014f8b_003dtest
mc014f8b_0041test
mc014f8b_0044test

MC requested 4.5.0

Request id	-	Run number	85899	Phys. GrP.	Top/Ewk	Gen type	Herwig
Events	246k	Xsect (pb)	4.652	Lum. (pb^-1)	52880	Req. St.	done
User	Vaijulis	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready
Physics type	tthar	Files	82	Events per file	3000	View tel	

Description text

inclusive tthar

Dataset definition has already been used to create actual datasets, so you cannot edit it. You may, however, add keywords to help you find it later. For more information, see the Datasets created using this definition.

Definition Name: mtest-to-top

Definition ID: 339

Date: 07/29/2002 14:35:02

Keyword Group: top

Owner: stdenis

Description: inclusive ttbar test

Dimension Query

LE_NAME mctest and RUN_NUMBER 85899



Definition ID: 339
Date: 07/29/2002 14:35:02
Keyword Group: top
Owner: stdenis
Description: inclusive ttbar test
Dimension Query
LE_NAME mctest and RUN_NUMBER 85899

Total File Count: 82 Total Event Count: 246000 Avg File Size: 413383

mc014f8b.004etest
mc014f8b.0043test
mc014f8b.0045test
mc014f8b.003etest
mc014f8b.003ttest
mc014f8b.0040test
mc014f8b.003dtest
mc014f8b.0041test
mc014f8b.0044test

Translate Constraints Results

[Clone](#) [Create Dataset](#)

Keyword Usages [?](#)

[Save All](#)

Definition Name: mtestc-top

Definition Id:

339

Create Date: 07/29/2002 14:35:02

top

Work Group:

Username:

stdenis

Description: inclusive ttbar test

Dimension Query
FILE_NAME mc%test and RUN_NUMBER 85899

Translate Constraints Results

Total File count:	Event count:	A
mc014f8b.004etest	82	
mc014f8b.0043test		
mc014f8b.0045test		
mc014f8b.003etest		
mc014f8b.003ftest		
mc014f8b.0040test		
mc014f8b.003dtest		
mc014f8b.0041test		
mc014f8b.0044test		

Dimensions

- A dimension is a nickname for a quantity that can be obtained from ANY part of the database
 - Examples: luminosity, run quality etc.
 - Definition (6wks) done with “sam add dimension” by sam support.
- For now, run numbers seem sufficient
 - Also, if stuff goes into the dfc as mctest, automatically in sam: so use this.

Request ID		Parameter		Value		Status		Actions	
Events	246k	Xsect (pb)	4.632	Lum. (pb^-1)	52880	Req. St.	done		
User	<u>Vainuus</u>	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready		
Physics type	tthar	Files	\$2	Events per file	3000		View tcl <th> </th> <th> </th>		
Description text									
inclusive tthar									

What is desired

- o select dataset for a process and with some parameters
- o discover no events or too few
- o pick up tcl and fill out a request form
- o submit request
- o receive notification that request is fulfilled and the name of the dataset.

Events	246k	Xsect (pb)	4.652	Lum. (pb^-1)	52880	Req. St.	done		
User	Vaitulis	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready		
Physics type	tbar	Files	\$2	Events per file	3000	View tcl			
Description text									
inclusive tbar									

What exists

- o Table to hold information on request
- o Indirect storage of datasets through DFC
- o ability to pick datasets based on dimensions and run AC++

What is missing

- o Entry Gui for requests – IS the DO stuff ok?
- o API that writes request to database – DO stuff?
- o Direct storage into sam with metadata – 1 mo.
- o Easy definition of dimensions (parameters)

Parameter	Description	Value	Unit	Min	Max	Step	Req. St.	Current St.	Notes
Events	Xsect (pb)	4.632	Lum. (pb^-1)	52880			Req. St.	done	
User	Vacuum	Exe. name	runMC	Exe. version	4.5.0		Exe. status	ready	
Physics type	tthar	Files	\$2	Events per file	3000			View tcl	

Description text

inclusive tthar

What is missing

- o Entry Gui for requests – IS the D0 stuff ok or limited?
- o API that writes request to database – D0 stuff?
can at least get information on db and get a browser page to check requests.
- o Direct storage into sam with metadata – underway.
- o Easy definition of dimensions (parameters) – 6 wks.

Events	246k	Xsect (pb)	4.652	Lum. (pb^-1)	52880	Req. St.	done
User	Vaitulis	Exe. name	runMC	Exe. version	4.5.0	Exe. status	ready
Physics type	tthar	Files	\$2	Events per file	3000		View tcl

Description text

inclusive tthar

Request

- o Request ID
- o work group
- o person
- o no. events
- o comments
- o priority
- o email
- o job name
- o status

Request Detail

- o request detail id
- o request id
- o application family
- o status
- o processing order

Request status: approved, complete
 finished,hold,new,partial,pending,
 running,terminated.

MC tools from DO

- Request system: sam create request
-dict=GetRequestKeywordDict.py
-group=dzero -num-events=100
-comment="Test job"
- Metadata system (keyword,value) like
topmass 170.

Status of tools

- Submission: in dev
- Modify request: in dev, mostly working
- Modify status: working
- Auto status mods: does not exist
- Request list: does not exist
- Get request: in development
- Keyword addition: being tested

Additional tasks

- Add physics parameters that are desired: store with file
- Put TCL portion of schema into production (can browse this now)
- Add request parameters: mc generator, random number, tcl collection id, gen xsect, validation status, physics type.
- Add dimensions for these parameters.
- Program to fetch params from CLOB into param for a file
- Add data run as a parameter in the request and the files as well.

Who

- Parameters and dimensions: Randy, Matt, etc (SBIR)
- GUI – Iain Bertram?
- Enter parameters and try: Rick, Frank, Dave
- Clob to param: Stefan (after batch adapter)

Conclusions

- Physics groups need to get the DB access under control and use present procedure for valid sets and make used sets
- Currently can make datasets based on things entered via the DFC and on run numbers and run AC++ from these files.
- Development underway to query datasets with richer choice of parameter
- Development underway but date unclear for Gui submission.